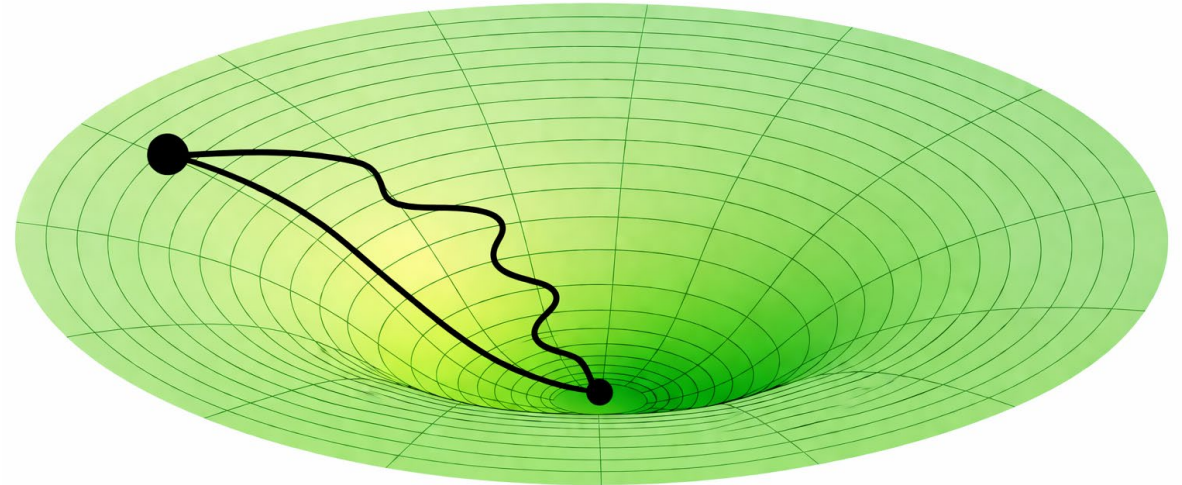
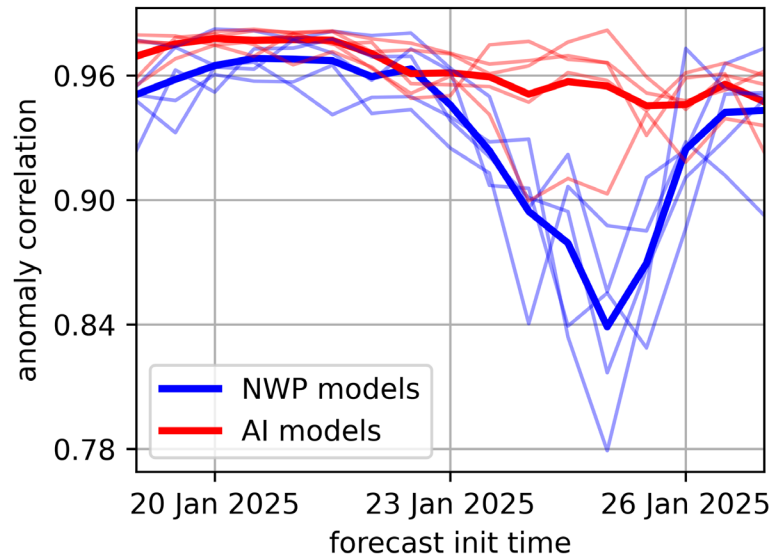


The physics of AI weather models

George Craig
Tobias Selz
Matthias Beylich
Kirsten Tempest

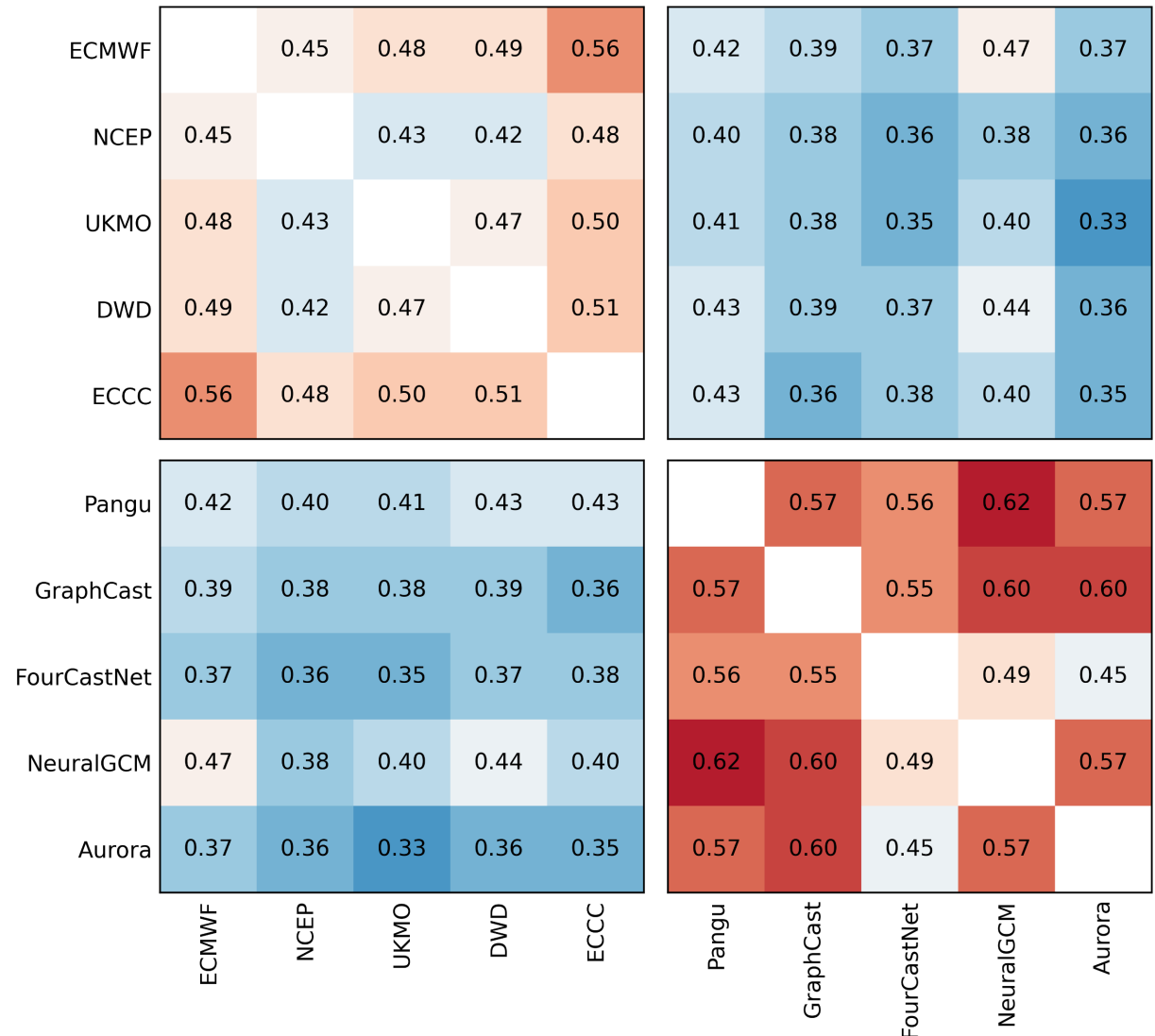


Are AI weather models different?



- Correlation of global RMS error for 500 hPa geopotential forecasts at day 5
- AI models seem more similar to each other than to NWP models
- Are AI models doing the same thing? A different description of the physics?

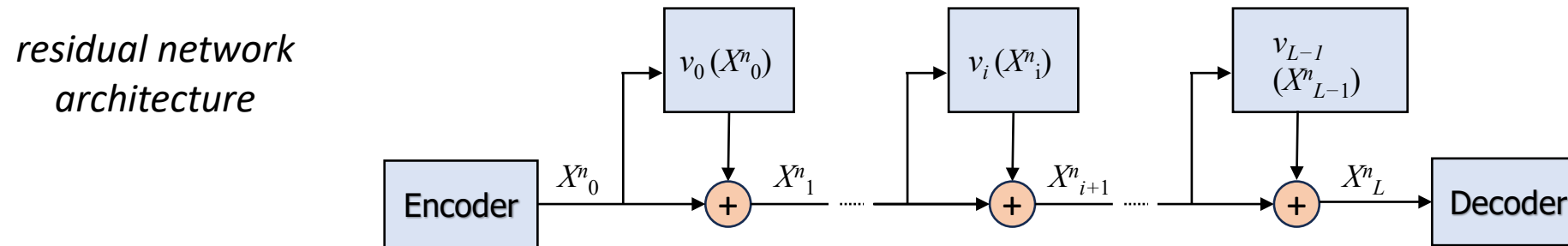
Correlation matrix, day 5.0



1. Are AI weather models differential equations?

- Deterministic model, or one denoising step of a diffusion model
- Initial conditions encoded to latent space vector X_0^n at each mesh point $n = 1 \dots N$
- Each processor layer $i = 1 \dots L$ reads the latent space vector and computes an increment

$$X_{i+1}^n = X_i^n + v_i(X_i^n)$$

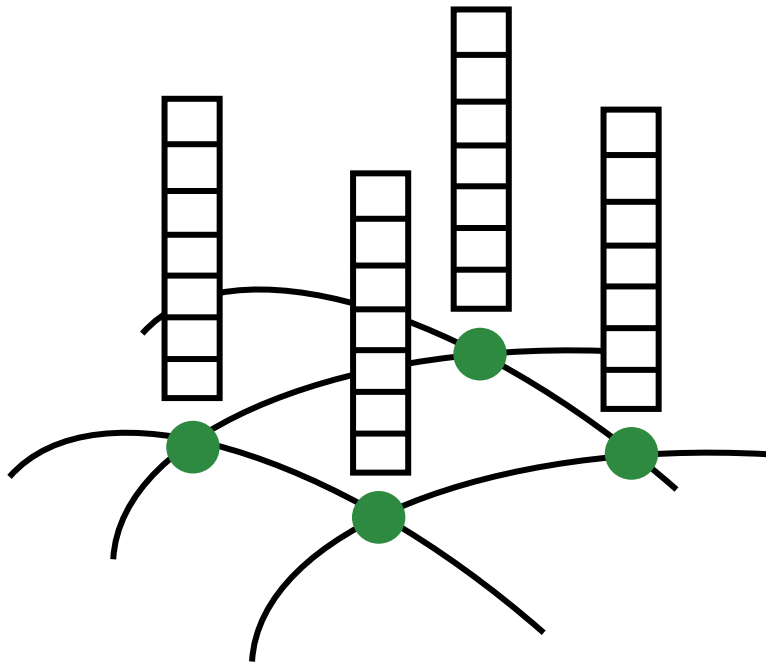


Continuum limit: Number of layers $L \rightarrow \infty$ (time-like variable s)

$$dX/ds = v(X^1(s), \dots, X^N(s), s)$$

- Coupled ordinary differential equations Neural ODE (*Chen et al. 2018, Marion et al. 2024*)

2. Lagrangian particle dynamics



- All properties of mesh point encoded in X^n
- Neural ODE describes movement of particle in configuration space

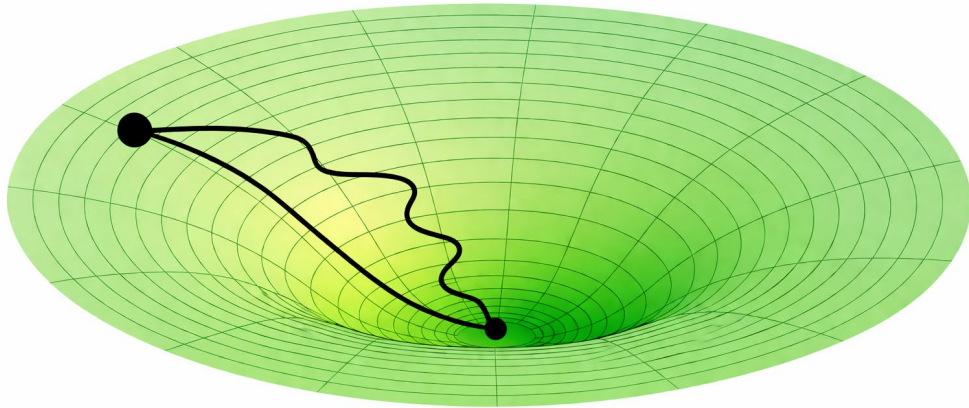
$$dX/ds = v(X^1(s), \dots, X^N(s), s)$$

- Equation has long-range interactions and shared weights (collisionless plasma, not ideal gas)
- McKean-Vlasov dynamics – rigorous for certain toy models (*Sander et al. 2022, Geshkovski et al. 2024*)

Continuum limit: number of particles $N \rightarrow \infty$ ($\Delta x \rightarrow 0$)

- Particles sample distribution with empirical measure $\alpha = \sum_n \delta(X - X^n)$
- Continuity equation $d\alpha/ds + \nabla \cdot (\alpha v) = 0$

3. Gradient flow



- Regularization/symmetries \Rightarrow bias for “direct” path from initial to final state (not rigorous!)
- Optimal transport: *Jordan-Kinderlehrer-Otto*
- Gradient flow: $v = \nabla_W(\mathcal{F})$
- Flow towards minimum of a “free energy”

$$\mathcal{F} = \int W(X, X') d\alpha d\alpha' + \int V(X) d\alpha + \int \alpha \log \alpha d\alpha$$

Continuum limit: $\Delta t \rightarrow 0$

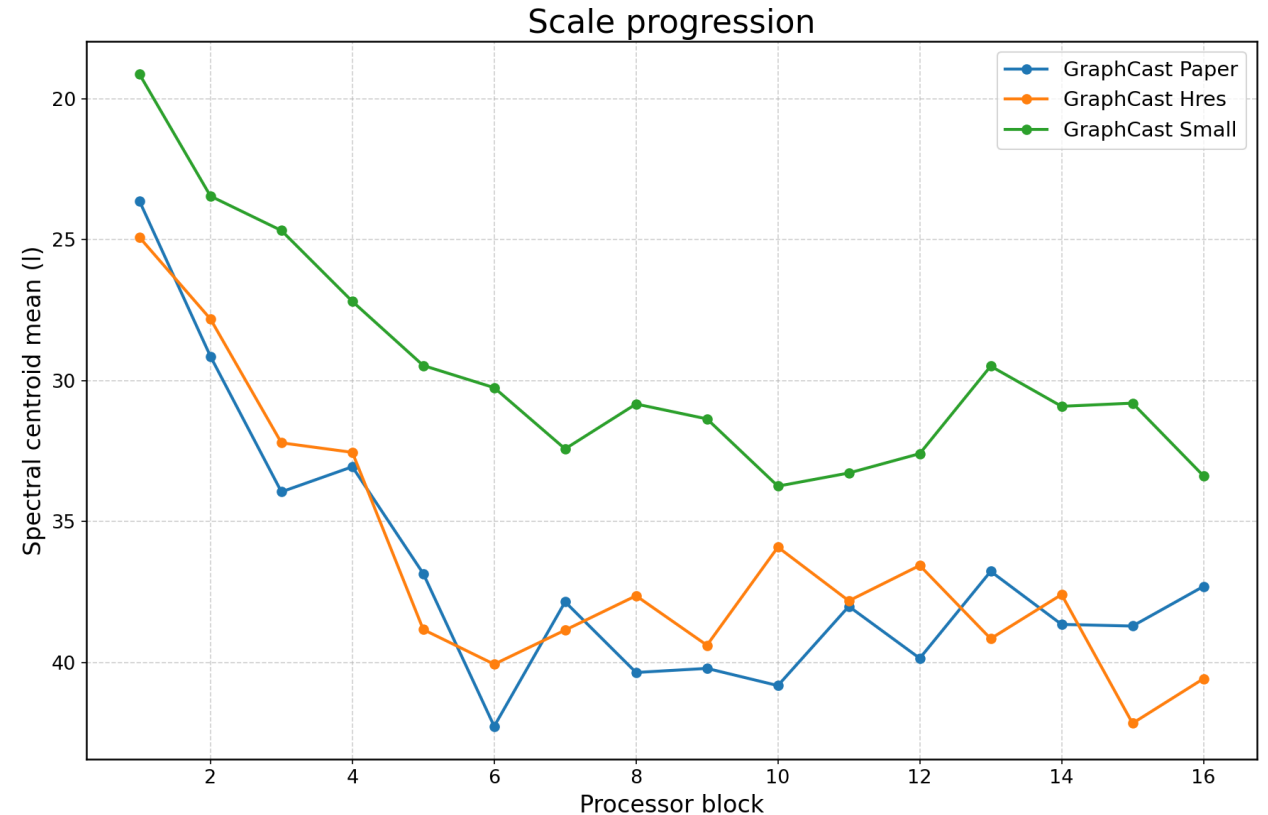
- System moves through equilibrium states (reversible thermodynamics)

Generative models: denoising/score matching gives a Wasserstein gradient flow

- Gradient flow with short step s , and long step σ (noise level)

Gradient flow: Large scales first

- Gradient flow suggests velocity decreases as equilibrium approached
- Hypothesis: Largest increments correspond to largest spatial scales
- Measure characteristic scale in latent space with spectral centroid (variance-weighted average wavenumber)
- Progression from large to small seen in three versions of GraphCast



Do two models have same internal representation?

Do they consider the same pairs of particles to be similar?

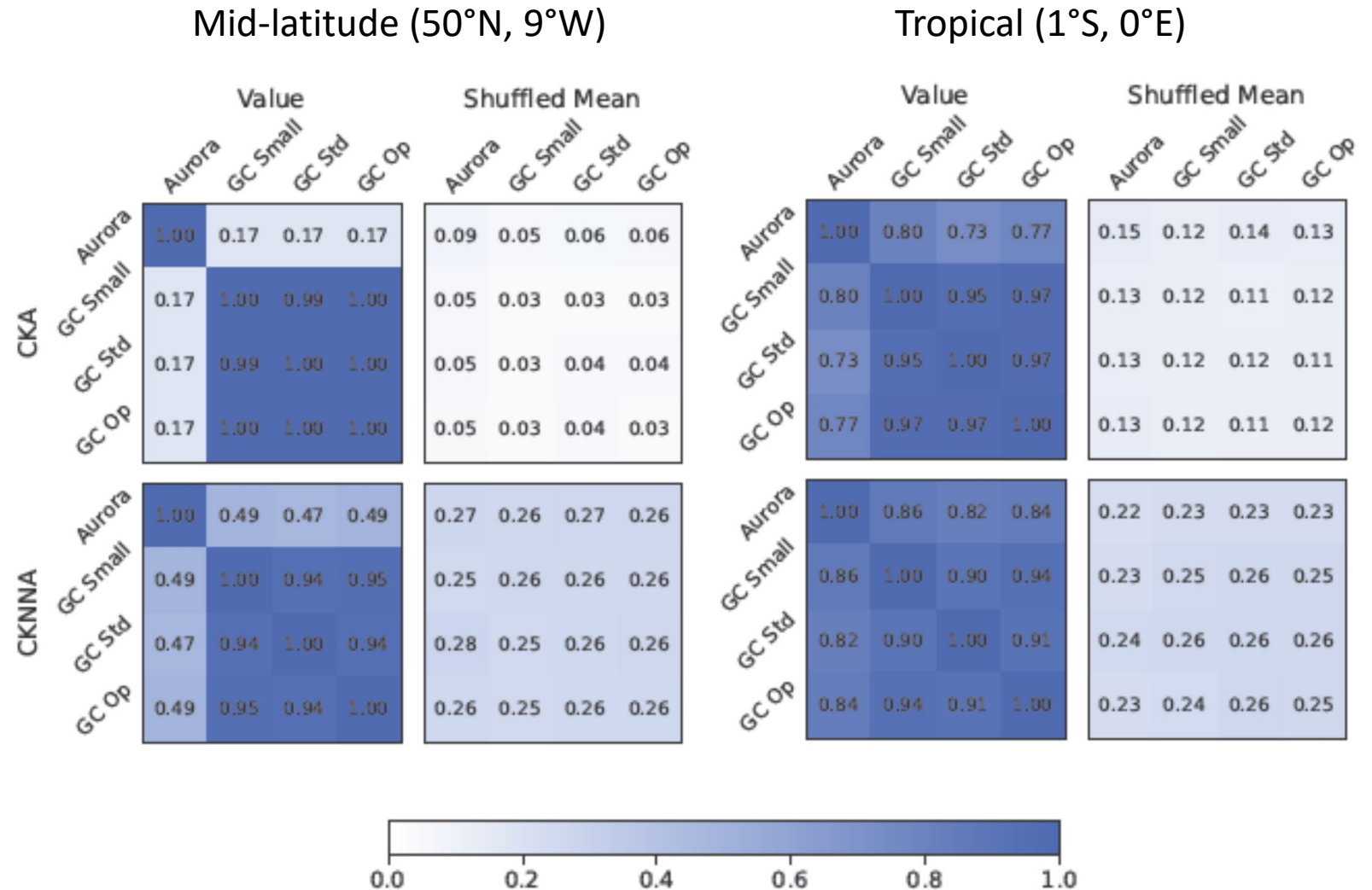
- For a single model, inner product $X_i^n \cdot X_i^m$ measures similarity between two particles
- **CKA – Centered Kernel Alignment** : Measure correlation between inner product similarity computed by the two models (“normalized Hilbert-Schmidt Independence Criterion for centered Gram matrices”)
- Problem: Most vectors in a high-dimensional space approx. orthogonal
- **CKNNA – Centered Kernel Nearest Neighbor Alignment**: Compare only nearby points

Typical results

- Molecular structure models (*Edamadaka et al. 2025*):
CKNNA = 0.03 – 0.55 (up to 0.85 for similar architectures)
- Language vs. vision models (*Huh et al. 2024*): CKNNA \approx 0.06 – 0.22 (significantly > 0)
- Stronger alignment as model capacity increases

GraphCast vs. Aurora (preliminary results)

- $i = L$ (equilibrium)
- Single locations, January 2020
- CKNNA for 10 neighbors
- GraphCast versions highly aligned (CKNNA ≈ 0.95)
- Aurora and GraphCast also aligned:
- Weakest alignment in winter mid-latitudes (≈ 0.5)
- Strongest in tropics (≈ 0.8)



Conclusions (tentative)

- AI weather models seem to share the same underlying description of reality: correlated forecast skill and latent space alignment (CKNNA)
- If this description approximates to physical laws, they likely have the form:
 1. Lagrangian particle system
 2. McKean-Vlasov dynamics, i.e. collisionless gas with long-range interactions
 3. Optimized to induce gradient flow towards minimum of a free energy
- Open questions
 1. How to interpret latent space coordinates – emergent phenomena?
 2. How does free energy depend on these coordinates?